

Stochastic Variational Partitioned Runge-Kutta Integrators for Constrained Systems

Nawaf Bou-Rabee* Houman Owhadi*

December 18, 2013

Abstract

Stochastic variational integrators for constrained, stochastic mechanical systems are developed in this paper. The main results of the paper are twofold: an equivalence is established between a stochastic Hamilton-Pontryagin (HP) principle in generalized coordinates and constrained coordinates via Lagrange multipliers, and variational partitioned Runge-Kutta (VPRK) integrators are extended to this class of systems. Among these integrators are first and second-order strongly convergent RATTLE-type integrators. We prove strong order of accuracy of the methods provided. The paper also reviews the deterministic treatment of VPRK integrators from the HP viewpoint.

1 Introduction

Since the foundational work of Bismut [1981], the field of stochastic geometric mechanics is emerging in response to the demand for tools to analyze the structure of continuous and discrete mechanical systems with uncertainty [9; 23; 5; 1; 17; 18; 12; 11; 15]. Within this context the goal of this paper is to develop efficient, structure-preserving integrators for long-time simulations of constrained, mechanical systems perturbed by white-noise forces and torques. Our strategy is to employ stochastic variational integrators (SVIs) [4].

Variational integration theory derives integrators for mechanical systems from discrete variational principles [24; 14; 25; 16]. The theory includes discrete analogs of the Lagrangian, Noether's theorem, the Euler-Lagrange equations, and the Legendre transform. Variational integrators can readily incorporate holonomic constraints (e.g., via Lagrange multipliers) and non-conservative effects (via their virtual work) [25; 16]. Altogether, this description of mechanics stands as a self-contained theory of mechanics akin to Hamiltonian, Lagrangian or Newtonian mechanics.

*Applied & Computational Mathematics (ACM), Caltech, Pasadena, CA 91125 (nawaf@acm.caltech.edu, owahadi@acm.caltech.edu).

Variational integrators are *symplectic*, i.e., the discrete flow map they define exactly preserves the continuous symplectic 2-form. Using backward error analysis one can show that symplectic integrators applied to Hamiltonian systems nearly preserve the energy of the continuous mechanical system for exponentially long periods of time and that the modified equations are also Hamiltonian [7]. Variational integrators are also distinguished by their ability to compute statistical properties of mechanical systems, such as in computing Poincaré sections, the instantaneous temperature of a system, etc.

Stochastic variational integrators are an extension of variational integrators to so-called stochastic mechanical systems. These systems are simple mechanical systems subject to certain random perturbations, and have their origins in Bismut’s foundational work [1]. Bismut showed that the flow of these stochastic mechanical systems extremize an action integral whose domain is the space of semimartingales on configuration space. Bismut’s work was further enriched and generalized to manifolds by recent work [12; 11]. Lazaro-Cami and Ortega show that this general class of stochastic Hamiltonian systems on manifolds extremizes a stochastic action defined on the space of manifold-valued semimartingales [12]. Moreover, it has been shown that for a subclass of these systems, one can prove a converse, namely, a.s. a curve satisfies so-called stochastic Hamilton’s equations if and only if it extremizes a stochastic action [4].

With this variational principle, one can design SVIs [4]. Like their deterministic counterparts, these methods have the advantage that they are symplectic, and in the presence of symmetry, satisfy a discrete Noether’s theorem. Moreover, symplectic methods for stochastic mechanical systems on vector spaces have been shown to capture the correct energy behavior even in the presence of dissipation [17; 18]. In particular, the energy injected or dissipated by the symplectic integrator is not an artificial function of the timestep. Moreover, the energy behavior is accurately captured by the integrator.

These structure-preserving properties are quite crucial. Consider, for instance, simulation of a simple mechanical system subject to random forces with amplitude ϵ . Suppose further the unperturbed system preserves energy, momentum, and possesses a first integral. Consider simulating this system with a higher-order accurate method, a standard integrator with simultaneous projection onto energy, momentum, and first integral level sets, and a stochastic variational integrator. If $\epsilon \neq 0$, a stochastic (physical) perturbation in the energy, momentum, and first integrals will appear. However, it is not clear how to modify the projection-based method to correctly capture these stochastic perturbations when $\epsilon \neq 0$. Moreover, the higher-order accurate method requires a time-step smaller than the amplitude of the perturbation in order to accurately represent its effects. And even then, if the time-span of integration is long enough systematic, artificial drift in these quantities will appear.

On the other hand, the stochastic perturbation in energy and momentum captured by the stochastic variational integrator is mechanical, i.e., it is only due to the ϵ -random forces. This is because these schemes define flows of discrete mechanical systems. Moreover, even when the amplitude of perturbation is not small, due to symplecticity we conjecture that SVIs on manifolds not

only possess good long-time energy behavior, but also perform well in computing statistical properties, such as autocorrelation functions and the empirical distribution. We will investigate these questions in future work.

As far as we can tell, the extension of these structure-preserving integrators to stochastic mechanical systems with holonomic constraints has not been completed. The main goal of this paper is to extend SVIs to holonomic constraints, and in particular, introduce ***constrained, stochastic variational partitioned Runge-Kutta*** (VPRK) methods for such systems. Within this family we exhibit a first-order strongly convergent, symplectic integrator for constrained mechanical systems. We use a technique due to Vanden-Eijnden and Cicotti [2006] to prove order of accuracy of the integrators that appear in this paper [23].

Future work will consider extensions of these schemes to Langevin equations with holonomic constraints. Continuous Langevin processes have been generalized to submanifolds $\Sigma \subset \mathbb{R}^n$ and shown to be ergodic [5]. The generalization to constraint submanifolds was done by appending holonomic constraints to Langevin equations. One can then check that the infinitesimal generator of the constrained Langevin process commutes with the Gibbsian density restricted to Σ to determine that the restricted Gibbsian measure is an invariant measure of the constrained Langevin process. To prove this measure is unique one uses standard arguments based on nondegeneracy of the diffusion and drift vector fields on the momentums. On this note it would be interesting to ascertain ergodicity of SVIs for constrained Langevin systems.

2 Constrained VPRK Integrators

This section is provided to fix notation and clarify some aspects of deterministic constrained VPRK integrators which will be relevant in the stochastic context. We also provide a novel proof showing that constrained VPRK integrators can be derived directly from a discrete variational principle without explicitly introducing a discrete Legendre transform. For more details on unconstrained VPRK integrators we refer the reader to [2].

The setting of this section is a real, n -dimensional vector space Q and a mechanical system with smooth, holonomic constraint function, $g : Q \rightarrow \mathbb{R}^k$, $k < n$, that has a regular value at $0 \in \mathbb{R}^k$. The mechanical system's configuration space is given by the *constraint submanifold*: $S = g^{-1}(0) \subset Q$. We will introduce Lagrange multipliers to prove an equivalence between a Hamilton-Pontryagin (HP) variational principle on S and a constrained HP variational principle on Q . We will then show how to discretize this system to obtain variational RATTLE integrators for constrained mechanical systems.

The variational and symplectic character of VPRK integrators is discussed in [22; 16; 7]. In what follows we will explicitly use the HP perspective, and specifically, extend the results in Hairer et al. [2006] to constrained systems using the HP perspective. It is worth mentioning, the Hamiltonian counterparts of the constrained VPRK methods, so-called symplectic partitioned Runge-Kutta

methods, are also well understood [10; 20; 6].

Discretization of HP Action We will adopt a HP viewpoint to describe the action integral of this mechanical system with constraints. The HP description unifies the Hamiltonian and Lagrangian descriptions of a mechanical system [26; 27; 2; 3].

Definition 2.1. *The **Pontryagin bundle** of a manifold M is defined as $PM = TM \oplus T^*M$. Fixing the interval $[a, b] \subset \mathbb{R}$ and $x_1, x_2 \in M$, define the **HP path space** on PM as*

$$\mathcal{C}(PM, x_1, x_2) = \{(q, v, p) \in C^\infty([a, b], PM) \mid q(a) = x_1, q(b) = x_2\}.$$

The HP path space is a smooth infinite-dimensional manifold. One can show that its tangent space at $c = (q, v, p) \in \mathcal{C}(PQ, x_1, x_2)$ consist of maps $w = (q, v, p, \delta q, \delta v, \delta p) \in C^\infty([a, b], T(PQ))$ such that $\delta q(a) = \delta q(b) = 0$.

Definition 2.2. *Fix $q_1, q_2 \in N$. Define the **unconstrained HP action** $\mathfrak{G} : \mathcal{C}(PQ, q_1, q_2) \rightarrow \mathbb{R}$ as*

$$\mathfrak{G} = \int_a^b \left[L(q, v) dt + \left\langle p, \frac{dq}{dt} - v \right\rangle dt \right].$$

To discretize \mathfrak{G} we first discretize the kinematic constraint: $dq/dt = v$. An s-stage RK method is employed to discretize the kinematic constraint. Let $[a, b]$ and N be given and define the fixed step size $h = (b - a)/N$ and $t_k = hk$, $k = 0, \dots, N$. The reason for using an s-stage RK discretization of the kinematic constraint is that the theory on such methods (order conditions, stability, and implementation) is mature. See, for instance, [8].

Definition 2.3. *Consider the first order differential equation*

$$\dot{q} = f(t, q), \quad q(0) = q_0, \quad q(t) \in Q. \quad (1)$$

*Let $b_i, a_{ij} \in \mathbb{R}$ ($i, j = 1, \dots, s$) and let $c_i = \sum_{j=1}^s a_{ij}$. An **s-stage RK** approximation is given by*

$$\begin{cases} Q_k^i &= q_k + h \sum_{j=1}^s a_{ij} f(t_k + c_j h, Q_k^j), \quad i = 1, \dots, s, \\ q_{k+1} &= q_k + h \sum_{j=1}^s b_j f(t_k + c_j h, Q_k^j). \end{cases} \quad (2)$$

The vectors q_k and Q_k^i are called external and internal stage vectors, respectively.

It follows that an s-stage RK method is fully determined by its a -matrix and b -vector which are typically displayed using the so-called Butcher tableau:

$$\begin{array}{c|ccc} c_1 & a_{11} & \cdots & a_{1s} \\ \vdots & \vdots & & \vdots \\ c_s & a_{s1} & \cdots & a_{ss} \\ \hline & b_1 & \cdots & b_s \end{array}$$

Suppose that $v(t)$, $t \in [a, b]$, is given. Then an s-stage RK approximant applied to $\dot{q} = v(t)$ yields:

$$\begin{cases} Q_k^i &= q_k + h \sum_{j=1}^s a_{ij} v(t_k + c_j h), \quad i = 1, \dots, s, \\ q_{k+1} &= q_k + h \sum_{j=1}^s b_j v(t_k + c_j h), \quad k = 0, \dots, N. \end{cases} \quad (3)$$

In what follows $v(t_k + c_i h)$ for $i = 1, \dots, s$ will be introduced as internal stage unknowns and will be determined as a critical point of a discrete action sum.

VPRK Integrator for Constrained Systems The VPRK method will be derived from a discretization of the HP action integral in which the kinematic constraint over the kth-time step is replaced with its discrete approximant: (3), and the integral of the Lagrangian over the kth-time step is approximated by the following quadrature:

$$\int_{t_k}^{t_k+h} L(q, v) dt \approx h b_i L(Q_k^i, V_k^i).$$

The constraint $g(q) = 0$ is enforced for all internal stage positions $\{Q_k^i\}$ using Lagrange multipliers as follows.

Definition 2.4. Fix two points q_1 and q_2 on Q and define the **discrete VPRK path space** as:

$$\mathcal{C}_d = \{(q, p, \{Q^i, V^i, P^i\}_{i=1}^s, \{\Lambda^i\}_{i=1}^s)_d : \{t_k\}_{k=0}^N \rightarrow T^*Q \times (TQ \oplus T^*Q)^s \times (\mathbb{R}^k)^s \mid q(0) = q_1, q(t_N) = q_2\},$$

and the **discrete constrained VPRK action sum** $\mathfrak{G}_d : \mathcal{C}_d \rightarrow \mathbb{R}$ by:

$$\begin{aligned} \mathfrak{G}_d = \sum_{k=0}^{N-1} \sum_{i=1}^s h & \left[b_i L(Q_k^i, V_k^i) + \left\langle p_k^i, (Q_k^i - q_k)/h - \sum_{j=1}^s a_{ij} V_k^j \right\rangle \right. \\ & \left. + \left\langle p_{k+1}^i, (q_{k+1} - q_k)/h - \sum_{j=1}^s b_j V_k^j \right\rangle + b_i \langle \Lambda_k^i, g(Q_k^i) \rangle \right]. \end{aligned}$$

The following condition on the coefficients of the s-stage RK method will be important in obtaining a well-defined discrete update on phase space that also respects the constraints [10].

Condition 2.1. Consider an s -stage RK method with given b -vector and a -matrix. Assume $b_k \neq 0$ and set $\hat{a}_{kj} = b_j - a_{jk}b_j/b_k$. The coefficients of the s -stage RK method satisfy:

$$a_{1i} = 0, \quad a_{si} = b_i, \quad b_i \neq 0, \quad i = 1, \dots, s$$

$$\left(\sum_{k=1}^s a_{ik} \hat{a}_{kj} \right)_{i,j=2}^s \text{ invertible.}$$

Under this condition on the coefficients of the s -stage RK method, one can prove the following theorem.

Theorem 2.5. Given an s -stage RK method with b -vector and a -matrix that satisfy condition 2.1, a Lagrangian system with smooth Lagrangian $L : TQ \rightarrow \mathbb{R}$ such that $\partial^2 L / \partial v^2$ is invertible, and smooth holonomic constraint function $g : Q \rightarrow \mathbb{R}^k$. A discrete curve $c_d \in \mathcal{C}_d(q_1, q_2)$ satisfies the following VPRK method:

$$\left\{ \begin{array}{lcl} Q_k^i & = & q_k + h \sum_{j=1}^s a_{ij} V_k^j, \\ q_{k+1} & = & q_k + h \sum_{j=1}^s b_j V_k^j, \\ P_k^i & = & p_k + h \sum_{j=1}^s \left(b_j - \frac{b_j a_{ji}}{b_i} \right) \left(\frac{\partial L}{\partial q}(Q_k^j, V_k^j) + \frac{\partial g}{\partial q}(Q_k^j)^* \cdot \Lambda_k^j \right), \\ p_{k+1} & = & p_k + h \sum_{j=1}^s b_j \left(\frac{\partial L}{\partial q}(Q_k^j, V_k^j) + \frac{\partial g}{\partial q}(Q_k^j)^* \cdot \Lambda_k^j \right), \\ P_k^i & = & \frac{\partial L}{\partial v}(Q_k^i, V_k^i), \\ g(Q_k^i) & = & 0. \end{array} \right. \quad (4)$$

for $i = 1, \dots, s$ and $k = 1, \dots, N-1$, if and only if it is a critical point of the function $\mathfrak{G}_d : \mathcal{C}_d \rightarrow \mathbb{R}$, that is, $\mathbf{d}\mathfrak{G}_d(c_d) = 0$. Moreover, there exist $\{\Lambda_k^i\}_{i=1}^s$ such that the discrete flow map defined by the above scheme, $F_h : T^*S \rightarrow T^*S$, preserves the canonical symplectic form on T^*S .

Proof. Under the assumptions of the theorem, the existence of a numerical solution of (4) and a discrete flow map $F_h : T^*S \rightarrow T^*S$ is guaranteed. That is, given $(q_k, p_k) \in T^*S$ one can solve (4) for $(q_{k+1}, p_{k+1}) \in T^*S$. In particular, the condition that $\partial^2 L / \partial v^2$ is invertible, implies one can eliminate $\{V_k^i\}_{i=1}^s$ using the Legendre transform of L . The condition 2.1 implies that one can determine $\{\Lambda_k^i\}_{i=1}^s$ so that $(q_{k+1}, p_{k+1}) \in T^*S$ [10].

The differential of $\mathfrak{G}_d(c_d)$ in the direction $z = (\{\delta q_k, \delta p_k\}, \{\delta Q_k^i, \delta V_k^i, p_k^i\}_{i=1}^s, \{\delta \Lambda_k^i\}_{i=1}^s)$

is given by:

$$\begin{aligned}
\mathbf{d}\mathfrak{G}_d \cdot z = & \sum_{k=0}^{N-1} \sum_{i=1}^s h b_i \left[\frac{\partial L}{\partial q}(Q_k^i, V_k^i) \cdot \delta Q_k^i + \frac{\partial L}{\partial v}(Q_k^i, V_k^i) \cdot \delta V_k^i \right] \\
& + h \left[\left\langle p_k^i, (\delta Q_k^i - \delta q_k)/h - \sum_{j=1}^s a_{ij} \delta V_k^j \right\rangle + \left\langle p_{k+1}, (\delta q_{k+1} - \delta q_k)/h - \sum_{j=1}^s b_j \delta V_k^j \right\rangle \right] \\
& + h \left[\left\langle \delta p_k^i, (Q_k^i - q_k)/h - \sum_{j=1}^s a_{ij} V_k^j \right\rangle + \left\langle \delta p_{k+1}, (q_{k+1} - q_k)/h - \sum_{j=1}^s b_j V_k^j \right\rangle \right] \\
& + h b_i \left[\langle \delta \Lambda_k^i, g(Q_k^i) \rangle + \left\langle \Lambda_k^i, \frac{\partial g}{\partial q}(Q_k^i) \cdot \delta Q_k^i \right\rangle \right].
\end{aligned}$$

Collecting terms with the same variations and summation by parts using the boundary conditions $\delta q_0 = \delta q_N = 0$ gives,

$$\begin{aligned}
\mathbf{d}\mathfrak{G}_d \cdot z = & \sum_{k=1}^{N-1} \sum_{i=1}^s \left(h b_i \frac{\partial g}{\partial q}(Q_k^i)^* \Lambda_k^i + h b_i \frac{\partial L}{\partial q}(Q_k^i, V_k^i) + p_k^i \right) \cdot \delta Q_k^i \\
& + \left(-p_{k+1} + p_k - \sum_{i=1}^s p_k^i \right) \cdot \delta q_k + h \left(b_i \frac{\partial L}{\partial v}(Q_k^i, V_k^i) - \sum_{j=1}^s a_{ji} p_k^j - b_i p_{k+1} \right) \cdot \delta V_k^i \\
& + h \left\langle \delta p_k^i, (Q_k^i - q_k)/h - \sum_{j=1}^s a_{ij} V_k^j \right\rangle + h \left\langle \delta p_{k+1}, (q_{k+1} - q_k)/h - \sum_{j=1}^s b_j V_k^j \right\rangle \\
& + h b_i \langle \delta \Lambda_k^i, g(Q_k^i) \rangle.
\end{aligned}$$

Since $\mathbf{d}\mathfrak{G}_d(c_d) = 0$ if and only if $\mathbf{d}\mathfrak{G}_d \cdot z = 0$ for all $z \in T_{c_d} \mathcal{C}_d$, one arrives at the desired equations with the elimination of p_k^i and the introduction of the internal stage variables $P_k^i = \partial L / \partial v(Q_k^i, V_k^i)$ for $i = 1, \dots, s$. Conversely, if c_d satisfies (4) then $\mathbf{d}\mathfrak{G}_d(c_d) = 0$.

We will employ the variational proof of symplecticity to prove that F_h is symplectic. This proof is standard, however, we provide it here to emphasize that the symplectic form that is exactly preserved by the method is the canonical one on T^*S .

Consider the subset of \mathcal{C}_d given by solutions of (4). Let $\hat{\mathfrak{G}}_d$ denote the restriction of \mathfrak{G}_d to this space. Since each of these solutions is determined by an initial point on T^*S , one can identify this space with T^*S , and hence, $\hat{\mathfrak{G}}_d : T^*S \rightarrow \mathbb{R}$. Since $\hat{\mathfrak{G}}_d$ is restricted to solution space,

$$\mathbf{d}\hat{\mathfrak{G}}_d(q_0, p_0) \cdot (\delta q_0, \delta p_0) = \langle p_N, \delta q_N \rangle - \langle p_0, \delta q_0 \rangle.$$

Observe that these boundary terms are the canonical one-forms on T^*S evaluated at $k = 0$ and $k = N$. Preservation of the symplectic form follows from $\mathbf{d}^2 \hat{\mathfrak{G}}_d = 0$. \square

Variational RATTLE for Constrained Systems The variational RATTLE integrator is the Lagrangian analog of the RATTLE algorithm originally proposed as a constrained version of Verlet in [21]. It was shown to be symplectic in [13], and was extended to general constrained Hamiltonian systems by [10]. It is defined by the following two-stage RK discretization of the kinematic constraint (implicit trapezoidal rule),

$$\begin{array}{c|cc} & 0 & 0 \\ & 1/2 & 1/2 \\ \hline & 1/2 & 1/2 \end{array}$$

Given h and (q_k, p_k) , the method determines (q_{k+1}, p_{k+1}) and two Lagrange multipliers Λ_k^1, Λ_k^2 by solving the following system of equations,

$$\left\{ \begin{array}{lcl} q_{k+1} & = & q_k + \frac{h}{2} (V_k^1 + V_k^2), \\ P_k^1 & = & p_k + \frac{h}{2} \left(\frac{\partial L}{\partial q}(q_k, V_k^1) + \frac{\partial q}{\partial q}(q_k)^* \Lambda_k^1 \right), \\ p_{k+1} & = & P_k^1 + \frac{h}{2} \left(\frac{\partial L}{\partial q}(q_{k+1}, V_k^2) + \frac{\partial q}{\partial q}(q_{k+1})^* \Lambda_k^2 \right) \\ 0 & = & g(q_{k+1}), \\ 0 & = & \frac{\partial q}{\partial q}(q_{k+1}) \cdot v_{k+1}, \\ p_{k+1} & = & \frac{\partial L}{\partial v}(q_{k+1}, v_{k+1}), \\ P_k^1 & = & \frac{\partial L}{\partial v}(q_k, V_k^1) = \frac{\partial L}{\partial v}(q_{k+1}, V_k^2). \end{array} \right. \quad (5)$$

By theorem 2.5, variational RATTLE defines a symplectic scheme. Moreover, as is well-known it is second-order accurate.

Variational Euler for Constrained Systems One can relax the conditions assumed in theorem 2.5 on the coefficients of the s-stage RK method [6]. For instance, if $b_s = 0$ in the s-stage RK method one can still obtain a well-defined variational integrator. Consider as an example the following two-stage RK method,

$$\begin{array}{c|cc} & 0 & 0 \\ & 1 & 0 \\ \hline & 1 & 0 \end{array}$$

The corresponding variational integrator is given by:

$$\left\{ \begin{array}{lcl} q_{k+1} & = & q_k + h \hat{v}_{k+1}, \\ \hat{p}_{k+1} & = & p_k + h \left(\frac{\partial L}{\partial q}(q_k, \hat{v}_{k+1}) + \frac{\partial q}{\partial q}(q_k)^* \Lambda_k^1 \right) \\ 0 & = & g(q_{k+1}), \\ \hat{p}_{k+1} & = & \frac{\partial L}{\partial v}(q_k, \hat{v}_{k+1}). \end{array} \right. \quad (6)$$

However, the corresponding discrete flow does not satisfy the “hidden” velocity constraint, and hence, does not define a map from T^*S to T^*S . To satisfy the

hidden constraint a projection step $(q_{k+1}, \hat{p}_{k+1}) \mapsto (q_{k+1}, p_{k+1})$ is taken:

$$\begin{cases} p_{k+1} &= \hat{p}_{k+1} + h \frac{\partial g}{\partial q}(q_{k+1})^* \Lambda_k^2, \\ 0 &= \frac{\partial g}{\partial q}(q_{k+1}) \cdot v_{k+1}, \\ p_{k+1} &= \frac{\partial L}{\partial v}(q_{k+1}, v_{k+1}). \end{cases} \quad (7)$$

One can check that this projection step defines a symplectic map. Thus, the composite map $(q_k, p_k) \mapsto (q_{k+1}, \hat{p}_{k+1}) \mapsto (q_{k+1}, p_{k+1})$ is symplectic. It is the Lagrangian version of the constrained symplectic Euler method [7].

Another example relaxing the assumptions in theorem 2.5 is given by regarding the 2-stage RK method as being single-stage implicit Euler,

$$\begin{array}{c|c} & 1 \\ \hline & 1 \end{array}$$

The corresponding variational integrator is:

$$\begin{cases} q_{k+1} &= q_k + h \hat{v}_{k+1}, \\ \hat{p}_{k+1} &= p_k + h \left(\frac{\partial L}{\partial q}(q_k, v_k) + \frac{\partial g}{\partial q}(q_k)^* \Lambda_k^1 \right) \\ 0 &= g(q_{k+1}), \\ \hat{p}_{k+1} &= \frac{\partial L}{\partial v}(q_{k+1}, \hat{v}_{k+1}). \end{cases} \quad (8)$$

We conclude with a theorem summarizing the structure-preserving properties of these constrained variational Euler methods.

Theorem 2.6. *The composition of one step of (6) or (8) and the projection step (7), defines a symplectic integrator on T^*S . Moreover, these integrators are first-order accurate.*

3 Constrained, Stochastic Mechanical Systems

Constrained, Stochastic HP Principle The setting in this section is an n -manifold Q and a *stochastic mechanical system* with smooth, holonomic constraint function, $g : Q \rightarrow \mathbb{R}^k$, $k < n$, that has a regular value $0 \in \mathbb{R}^k$. Its configuration space is given by the *constraint submanifold*: $S = g^{-1}(0)$. In this section we will introduce Lagrange multipliers to prove an equivalence between a stochastic variational principle on S and a constrained, stochastic variational principle on Q .

We will adopt a HP viewpoint to describe this mechanical system with random perturbations and will refer to this perturbed system as a ***stochastic mechanical system***. The HP principle unifies the Hamiltonian and Lagrangian descriptions of a mechanical system [26; 27; 2; 3]. Roughly speaking, in the stochastic context it states the following critical point condition on PS (cf. definition 2.1),

$$\delta \int_a^b \left[L(q, v) dt + \sum_{i=1}^m \gamma_i(q) \circ dW_i + \left\langle p, \frac{dq}{dt} - v \right\rangle dt \right] = 0,$$

where $(q(t), v(t), p(t)) \in PS$ are varied arbitrarily and independently with end-point conditions $q(a)$ and $q(b)$ fixed, and $\gamma_i : Q \rightarrow \mathbb{R}$ and W_i for $i = 1, \dots, m$ are **stochastic potentials** and Wiener processes respectively. This principle builds in a Legendre transform, stochastic Hamilton's equations and stochastic Euler-Lagrange equations. The action integral in the above principle, consists of two Lebesgue integrals with respect to the Lebesgue measure dt and m Stratonovich stochastic integrals. This action is random, i.e., for every sample point $\omega \in \Omega$ one will obtain a different, time-dependent Lagrangian system. However, each system possesses a variational structure as made precise in [4]. The following definitions will be useful to state the constrained, variational principle of HP for mechanical systems with holonomic constraints.

Definition 3.1. Let $i : S \rightarrow Q$ be the inclusion mapping. Fixing the interval $[a, b]$, define **constrained HP path space** as

$$\mathcal{C}_c^{HP}([a, b], q_1, q_2) = \{(q, v, p) : [a, b] \rightarrow PS \mid \\ q \in C^1([a, b], S), (v, p) \in C^0([a, b]), q(a) = q_1, q(b) = q_2\},$$

and the **unconstrained HP path space** as

$$\mathcal{C}^{HP}([a, b], q_1, q_2) = \{(q, v, p) : [a, b] \rightarrow PQ \mid \\ q \in C^1([a, b], Q), (v, p) \in C^0([a, b]), q(a) = i(q_1), q(b) = i(q_2)\}.$$

Let (Ω, \mathcal{F}, P) be a probability space, $\mathcal{F}_t, t \in [a, b]$, be a nondecreasing family of σ -subalgebras of \mathcal{F} , $\omega \in \Omega$, and $(W_i(t, \omega), \mathcal{F}_t), i = 1, \dots, m$, be independent, real-valued Wiener processes. In terms of these Wiener processes, we define the following.

Definition 3.2. Set $L^S = L|_{TS}$ and $\gamma_j^S = \gamma_j|_S$ for $j = 1, \dots, m$. Moreover define the **unconstrained action** $\mathfrak{G} : \Omega \times \mathcal{C}([a, b], q_1, q_2) \rightarrow \mathbb{R}$ as

$$\mathfrak{G} = \int_a^b \left[L(q, v)dt + \sum_{j=1}^m \gamma_j(q) \circ dW_j + \left\langle p, \frac{dq}{dt} - v \right\rangle dt \right].$$

and the **constrained action** as $\mathfrak{G}_c = \mathfrak{G}|_{\mathcal{C}_c^{HP}}$.

The unconstrained HP path space is a smooth infinite-dimensional manifold. One can show that its tangent space at $c = (q, v, p) \in \mathcal{C}^{HP}([a, b], q_1, q_2)$ consist of maps $w = (q, v, p, \delta q, \delta v, \delta p) \in C^0([a, b], T(PQ))$ such that $\delta q(a) = \delta q(b) = 0$ and $q, \delta q$ are of class C^1 .

As opposed to using generalized coordinates on PS , we wish to describe the mechanical system using constrained coordinates on PQ and introduce Lagrange multipliers to enforce the constraint. However, because of the stochastic component of the action, the standard Lagrange multiplier theorem will not apply directly and one cannot introduce Lagrange multipliers in the standard way. Instead, we will introduce the Lagrange multiplier using the following.

Definition 3.3. Given $f_1 \in C^0([a, b], T^*Q)$ and $f_2 \in C^1([a, b], TQ)$, define

$$\int_a^t \langle df_1, f_2 \rangle := \langle f_1, f_2 \rangle|_a^t - \int_a^t \left\langle f_1, \frac{df_2}{dt} \right\rangle dt, \quad t \in [a, b].$$

Differentiability of the flow map on T^*S will be defined in the mean-squared sense. In the following we define mean-squared derivatives on a Banach space E with the understanding that this notion can be extended to any manifold using a local representative of the flow map.

Definition 3.4 (Mean-squared Derivative). *The mean squared norm of $f : E \times \Omega \rightarrow E$ is given by:*

$$\|f(x, \omega)\| = (\mathbb{E}(|f(x, \omega)|^2))^{1/2}$$

Using this norm one can define the derivative of f in the standard way, i.e., f is mean squared differentiable at $a \in E$ if there is a bounded linear map $Df(a, \omega) : E \rightarrow E$ that satisfies,

$$\lim_{\delta \rightarrow 0} \frac{\|f(a + \delta, \omega) - f(a, \omega) - Df(a, \omega) \cdot \delta\|}{\|\delta\|} \rightarrow 0.$$

With the above definitions we prove the following.

Theorem 3.5 (Constrained, Stochastic HP Principle). *Given a stochastic mechanical system with Lagrangian $L : TQ \rightarrow \mathbb{R}$ such that $\partial^2 L / \partial v^2$ is invertible, stochastic potentials $\gamma_j : Q \rightarrow \mathbb{R}$ for $j = 1, \dots, m$, and holonomic constraint $g : Q \rightarrow \mathbb{R}^k$ with $S = g^{-1}(0)$. Then the following are equivalent:*

- (i) $z = (q, v, p) \in \mathcal{C}_c^{HP}([a, b], q_1, q_2)$ extremizes \mathfrak{G}_c .
- (ii) $z = (q, v, p) \in \mathcal{C}_c^{HP}([a, b], q_1, q_2)$ satisfies stochastic HP equations

$$\begin{cases} dq &= v dt, \\ dp &= \frac{\partial L^S}{\partial q}(q, v) dt + \sum_{j=1}^m \frac{\partial \gamma_j^S}{\partial q}(q) \circ dW_j, \\ p &= \frac{\partial L^S}{\partial v}(q, v). \end{cases} \quad (9)$$

- (iii) There exists $\lambda \in C^0([a, b], \mathbb{R}^k)$ such that $z = (q, v, p) \in \mathcal{C}^{HP}([a, b], i(q_1), i(q_2))$ and λ extremize the augmented action $\bar{\mathfrak{G}}(z, \lambda) = \mathfrak{G}(z) + \langle \lambda, \Phi(z) \rangle$ where $\Phi(z)(t) = dg/dt(q(t))$ and $\langle \lambda, \Phi(z) \rangle = \int_a^b \langle \lambda(s), \Phi(z)(s) \rangle ds$.
- (iv) There exists $\lambda \in C^0([a, b], \mathbb{R}^k)$ such that $z = (q, v, p) \in \mathcal{C}^{HP}([a, b], i(q_1), i(q_2))$ and λ satisfy the constrained, stochastic HP equations

$$\begin{cases} dq &= v dt, \\ dp &= \frac{\partial L}{\partial q}(q, v) dt + \sum_{j=1}^m \frac{\partial \gamma_j}{\partial q}(q) \circ dW_j + \frac{\partial g}{\partial q}(q)^* \cdot d\lambda, \\ p &= \frac{\partial L}{\partial v}(q, v), \\ g(q) &= 0. \end{cases} \quad (10)$$

Moreover, the flows of (9) and (10) are mean-square symplectic.

Remark 3.1. The constrained, stochastic HP equations should be thought of in integral form. In particular, the Lagrange multiplier term in (10) by definition 3.3 satisfies:

$$\int_a^t \left\langle \frac{\partial g}{\partial q}(q)^* \cdot d\lambda, w \right\rangle = \left\langle \lambda, \frac{\partial g}{\partial q}(q)w \right\rangle \Big|_a^t - \int_a^t \left\langle \lambda, \frac{d}{dt} \frac{\partial g}{\partial q}(q)w \right\rangle dt,$$

for any $w \in \mathbb{R}^k$.

Proof. The stochastic HP principle states that (i) and (ii) are equivalent [4]. Assume that (iii) is true. Then for all $v_z = (v_q, v_v, v_p) \in T_z \mathcal{C}^{HP}$,

$$\mathbf{d}\bar{\mathfrak{G}}(z, \lambda) \cdot (v_z, v_\lambda) = \mathbf{d}\mathfrak{G}(z) \cdot v_z + \langle v_\lambda, \Phi(z) \rangle + \langle \lambda, \mathbf{d}\Phi(z) \cdot v_z \rangle = 0 \quad (11)$$

Since $\Phi(z) = 0$ for $z \in \mathcal{C}_c^{HP}$ and $T_{q(t)}g(q(t)) \cdot v_q(t) = 0$ for $v_q(t) \in T_q S$,

$$\mathbf{d}\bar{\mathfrak{G}}(z, \lambda) \cdot (v_z, v_\lambda) = \mathbf{d}\mathfrak{G}(z) \cdot v_z = 0, \quad \forall v_z \in T_z \mathcal{C}_c^{HP}$$

which implies (i) and (ii).

Moreover, expanding (11) and setting $(v_q, v_v, v_p) = (\delta q, \delta v, \delta p)$ and $v_\lambda = \delta \lambda$ yields,

$$\begin{aligned} \mathbf{d}\bar{\mathfrak{G}}(z, \lambda) \cdot (\delta q, \delta v, \delta p, \delta \lambda) &= \int_a^b \left[\left\langle \frac{\partial L}{\partial q}, \delta q \right\rangle ds + \sum_{i=1}^m \left\langle \frac{\partial \gamma_i}{\partial q}, \delta q \right\rangle \circ dW_i + \left\langle \frac{\partial L}{\partial v}, \delta v \right\rangle ds \right. \\ &\quad \left. + \left\langle \delta p, \frac{dq}{dt} - v \right\rangle ds + \left\langle p, \delta \frac{dq}{dt} - \delta v \right\rangle ds + \left\langle \delta \lambda, \frac{d}{dt} g(q) \right\rangle ds + \left\langle \lambda, \frac{d}{dt} \left(\frac{\partial g}{\partial q} \cdot \delta q \right) \right\rangle ds \right] = 0. \end{aligned}$$

Consider the terms involving δp , δv and $\delta \lambda$. Since these variations are arbitrary, the following hold:¹

$$\frac{dq}{dt} = v, \quad \frac{\partial L}{\partial v}(q, v) = p, \quad \frac{d}{dt} g(q) = 0.$$

Since $g(q(a)) = g(i(q_1)) = 0$, $d/dt(g(q)) = 0$ implies that $g(q(t)) = 0$ for all $t \in [a, b]$.

Collecting the variations with respect to δq in the differential gives,

$$\int_a^b \left[\frac{\partial L}{\partial q} \cdot \delta q ds + \left\langle p, \delta \frac{dq}{dt} \right\rangle ds + \sum_{j=1}^m \frac{\partial \gamma_j}{\partial q} \cdot \delta q \circ dW_j + \left\langle \lambda, \frac{d}{dt} \left(\frac{\partial g}{\partial q} \cdot \delta q \right) \right\rangle ds \right] = 0.$$

Using definition 3.3, we introduce the following function $I : \mathcal{C}^{HP}(q_1, q_2, [a, b]) \times C([a, b], \mathbb{R}^k) \times C^1([a, b], TQ) \rightarrow \mathbb{R}$,

$$I(q, v, p, \lambda, f) = \int_a^b \left[\left(\frac{\partial L}{\partial q} ds + \sum_{j=1}^m \frac{\partial \gamma_j}{\partial q} \circ dW_j - dp - \frac{\partial g}{\partial q}^* \cdot d\lambda \right) \cdot f \right].$$

¹This follows from the basic lemma that if $f, g \in C^0([a, b], \mathbb{R})$ and g is arbitrary then $\int_a^b f(t)g(t)dt = 0 \iff f(t) = 0 \quad \forall t \in [a, b]$.

In the following it is shown that if $I(q, v, p, \lambda, f) = 0$ for arbitrary f of class C^1 then $c = (q, v, p, \lambda)$ satisfies (10).

Let $\{U_\alpha, g_\alpha\}$ be a partition of unity on $PQ \times \mathbb{R}^k$. Expand I in terms of this partition of unity,

$$I = \sum_{\alpha} \int_a^b \left[g_{\alpha}(q, v, p, \lambda) \left(\frac{\partial L}{\partial q} dt + \sum_{j=1}^m \frac{\partial \gamma_j}{\partial q} \circ dW_j - dp - \frac{\partial g^*}{\partial q} \cdot d\lambda \right) \cdot f \right].$$

Since the curves $z = (q, v, p)$ and λ are compactly supported, only a finite number of the g_{α} are nonzero. For each g_{α} nonzero, the terms in the integral can be expressed in local coordinates. Observe that since $dq = v dt$, the Stratonovich-Ito conversion formula implies that,

$$\int_a^b g_{\alpha} \frac{\partial \gamma_j}{\partial q} \cdot \delta q \circ dW_j = \int_a^b g_{\alpha} \frac{\partial \gamma_j}{\partial q} \cdot \delta q dW_j$$

for $j = 1, \dots, m$.

We will select f to single out the β th-component of the covector field in I . Introduce the following function $h : \mathbb{R} \rightarrow \mathbb{R}$ for this purpose:

$$h(t) = 2\frac{t}{\epsilon} - \frac{t^2}{\epsilon^2}.$$

Observe that $h(0) = 0$, $h(\epsilon) = 1$, and $h'(\epsilon) = 0$. Let $\{e_{\beta}\}$ be a basis for the model space of Q . Now fix β , and define $f_{\epsilon} \in C^1([a, b], TQ)$ in local coordinates as:

$$f_{\epsilon}(s) = \begin{cases} h(s-a)e_{\beta} & \text{if } a \leq s \leq a+\epsilon, \\ e_{\beta} & \text{if } a+\epsilon < s < t-\epsilon, \\ h(t-s)e_{\beta} & \text{if } t-\epsilon \leq s \leq t, \\ 0 & \text{if } t < s \leq b. \end{cases}$$

Introduce the following label to simplify subsequent calculations,

$$A(s) = \left(\frac{\partial L}{\partial q}(q(s), v(s)) ds + \sum_{i=1}^m \frac{\partial \gamma_i}{\partial q}(q(s)) dW_i(s) - dp(s) - \frac{\partial g}{\partial q}(q(s))^* \cdot d\lambda(s) \right) \cdot e_{\beta}.$$

In terms of $A(s)$, one can write

$$I(q, v, p, \lambda, f_{\epsilon}) = \sum_{\alpha} \left[\int_a^{a+\epsilon} h(s-a)g_{\alpha}(s)A(s) + \int_{a+\epsilon}^{t-\epsilon} g_{\alpha}(s)A(s) + \int_{t-\epsilon}^t h(t-s)g_{\alpha}(s)A(s) \right].$$

We will show in the mean squared norm,

$$\lim_{\epsilon \rightarrow 0} I(q, v, p, \lambda, f_{\epsilon}) = \sum_{\alpha} \int_a^t g_{\alpha} A(s) =: I^*. \quad (12)$$

Using this result and the Borel-Cantelli lemma, one can deduce there exists $\{\epsilon_n\}$ that converges to 0 such that $I(q, v, p, \lambda, f_{\epsilon_n})$ a.s. converges to I^* . It follows that $I^* = 0$ almost surely.

We proceed to prove (12). Since $(a + b)^2 \leq 2a^2 + 2b^2$,

$$\begin{aligned} & \left\| \sum_{\alpha} \int_a^t g_{\alpha} A(s) - I(q, v, p, \lambda, f_{\epsilon}) \right\|^2 \\ &= \left\| \sum_{\alpha} \int_a^{a+\epsilon} (1 - h(s - a)) g_{\alpha} A(s) + \int_{t-\epsilon}^t (1 - h(t - s)) g_{\alpha} A(s) \right\|^2 \\ &\leq 2 \left\| \sum_{\alpha} \int_a^{a+\epsilon} (1 - h(s - a)) g_{\alpha} A(s) \right\|^2 + 2 \left\| \sum_{\alpha} \int_{t-\epsilon}^t (1 - h(t - s)) g_{\alpha} A(s) \right\|^2. \end{aligned}$$

We will only show how to bound the first term since bounding the second term is very similar. By continuity of (q, v, p, λ) , one can pick $\epsilon > 0$ small enough so that the support of (q, v, p, λ) lies in a single chart. Therefore,

$$\begin{aligned} & \left\| \sum_{\alpha} \int_a^{a+\epsilon} (1 - h(s - a)) g_{\alpha} A(s) \right\|^2 \\ &= \left\| \int_a^{a+\epsilon} (1 - h(s - a)) \left(\frac{\partial L}{\partial q} ds + \sum_{j=1}^m \frac{\partial \gamma_j}{\partial q} dW_j - dp - \frac{\partial g^*}{\partial q} \cdot d\lambda \right) \cdot e_{\beta} \right\|^2 \\ &\leq 4 \left\| \int_a^{a+\epsilon} (1 - h(s - a)) \frac{\partial L}{\partial q^{\beta}} ds \right\|^2 + 4 \sum_{j=1}^m \left\| \int_a^{a+\epsilon} (1 - h(s - a)) \frac{\partial \gamma_j}{\partial q^{\beta}} dW_j \right\|^2 \\ &\quad + 4 \left\| \int_a^{a+\epsilon} (1 - h(s - a)) dp \cdot e_{\beta} \right\|^2 + 4 \left\| \int_a^{a+\epsilon} (1 - h(s - a)) \left\langle \frac{\partial g^*}{\partial q} \cdot d\lambda, e_{\beta} \right\rangle \right\|^2. \end{aligned}$$

Since $\frac{\partial L}{\partial q^{\beta}}$ is continuous on $s \in [a, a + \epsilon]$, the first term can be bounded,

$$\left\| \int_a^{a+\epsilon} (1 - h(s - a)) \frac{\partial L}{\partial q^{\beta}} ds \right\|^2 \leq \frac{M^2 \epsilon^2}{9}.$$

Similarly, by the Ito isometry and since $\frac{\partial \gamma_j}{\partial q^{\beta}}$ is continuous on $s \in [a, a + \epsilon]$, the second m terms can similarly be bounded, e.g., the j th Stratonovich integral can be bounded as follows,

$$\left\| \int_a^{a+\epsilon} (1 - h(s - a)) \frac{\partial \gamma_j}{\partial q^{\beta}} dW_j \right\|^2 = E \left(\int_a^{a+\epsilon} \left| (1 - h(s - a)) \frac{\partial \gamma_j}{\partial q^{\beta}} \right|^2 ds \right) \leq \frac{M^2 \epsilon}{5}.$$

Using (3.3) and the integral mean value theorem, there exists a real constant $c_1 \in [0, 1]$ such that:

$$\begin{aligned} \left\| \int_a^{a+\epsilon} (1 - h(s - a)) dp \cdot e_{\beta} \right\|^2 &= \left\| (1 - h(s - a)) p_{\beta}(s) \Big|_a^{a+\epsilon} + \int_a^{a+\epsilon} p_{\beta}(s) h'(s - a) ds \right\|^2 \\ &= \left\| -p_{\beta}(a) + p_{\beta}(a + c_1 \epsilon) \right\|^2. \end{aligned}$$

Similarly, there exist constants $c_2, c_3 \in [0, 1]$ such that

$$\begin{aligned}
& \left\| \int_a^{a+\epsilon} (1 - h(s-a)) \left\langle \frac{\partial g^*}{\partial q} \cdot d\lambda, e_\beta \right\rangle \right\|^2 = \\
& = \left\| (1 - h(s-a)) \left\langle \frac{\partial g^*}{\partial q} \cdot \lambda, e_\beta \right\rangle \Big|_a^{a+\epsilon} - \int_a^{a+\epsilon} (1 - h(s-a)) \lambda_\beta \frac{d}{ds} \frac{\partial g}{\partial q^\beta}(q(s)) ds \right. \\
& \quad \left. + \int_a^{a+\epsilon} \left\langle \lambda, \frac{\partial g}{\partial q} \cdot e_\beta \right\rangle h'(s-a) ds \right\|^2 \\
& = \left\| - \left\langle \frac{\partial g}{\partial q}(q(a))^* \cdot \lambda(a), e_\beta \right\rangle - \frac{\epsilon}{3} \lambda_\beta(a + c_2\epsilon) \frac{\partial g}{\partial q^\beta}(q(a + c_2\epsilon)) \right. \\
& \quad \left. + \left\langle \frac{\partial g}{\partial q}(q(a + c_3\epsilon))^* \cdot \lambda(a + c_3\epsilon), e_\beta \right\rangle \right\|^2.
\end{aligned}$$

Since p_β and λ are of class C^0 and g is smooth, as $\epsilon \rightarrow 0$ these terms vanish. Since β is arbitrary we have proved (12). Therefore, almost surely: if $c = (z, \lambda)$ is a critical point of \mathfrak{G} then c satisfies (10).

On the other hand, assume (iv) is true. Then almost surely: if $c = (z, \lambda)$ satisfies the constrained, stochastic HP equations, then it is a critical point of \mathfrak{G} . This direction is easy to confirm, since as a solution to the constrained, stochastic HP equations c is a measurable diffusion process. In fact, this direction is similar to the one Bismut originally established, namely that the solution of stochastic Hamilton's equations extremize an action function; albeit a different stochastic action is used in this proof [1].

Assume that (i) is true, i.e.,

$$\mathbf{d}\mathfrak{G}_c(z) \cdot v_z = 0 \quad \forall \quad v_z \in T_z \mathcal{C}_c^{HP}.$$

Define $\mathcal{C}^Q = \{q \in C^1([a, b], Q) \mid q(a) = i(q_1), q(b) = i(q_2)\}$. Let $\Psi(q)(t) = g(q(t))$ where $\Psi : \mathcal{C}^Q \rightarrow C^1([a, b], \mathbb{R}^k)$. By the local onto theorem, there exist charts (U, φ) of \mathcal{C}^Q and an open set $V \subset \mathbb{R}^k$ such that $\Psi \circ \varphi^{-1} : U' \times V \rightarrow V$ is a projection onto the second factor and $\varphi(q) = (x, y) \in U' \times V$ for $q \in U$. Set $v_q = T_{x,y} \varphi^{-1}(v_x, v_y)$ where $v_x \in U'$ and $v_y \in V$. Set $\tilde{\mathfrak{G}}(q) = \mathfrak{G}(q, v, p)$. Then,

$$\begin{aligned}
\mathbf{d}\tilde{\mathfrak{G}}(z, \lambda) \cdot (v_z, v_\lambda) &= T\tilde{\mathfrak{G}}(q) \cdot v_q + \left\langle \lambda, T\Psi(q) \cdot \frac{d}{dt}(v_q) \right\rangle \\
&= T_2(\tilde{\mathfrak{G}} \circ \varphi^{-1})(x, y) \cdot v_y + \left\langle \lambda, \frac{d}{dt}(v_y) \right\rangle.
\end{aligned}$$

However, there exists λ such that,

$$T_2(\tilde{\mathfrak{G}} \circ \varphi^{-1})(x, y) \cdot v_y + \left\langle \lambda, \frac{d}{dt}(v_y) \right\rangle = 0$$

for all $v_y \in V$. To see this expand this expression,

$$\int_a^b \left[\left\langle \frac{\partial L}{\partial y}(\varphi^{-1}(x, y), v), v_y \right\rangle dt + \sum_{j=1}^m \left\langle \frac{\partial \gamma_j}{\partial y}(\varphi^{-1}(x, y)), v_y \right\rangle \circ dW_j + \left\langle \lambda, \frac{d}{dt} v_y \right\rangle dt \right] = 0$$

and note that a solution is given by:

$$d\lambda = \frac{\partial L}{\partial y}(\varphi^{-1}(x, y), v)dt + \sum_{j=1}^m \frac{\partial \gamma_j}{\partial y}(\varphi^{-1}(x, y)) \circ dW_j$$

which is an SDE for λ , with globally Lipschitz drift and diffusion vector fields. Hence, there exists λ such that,

$$\mathbf{d}\bar{\Theta}(z, \lambda) \cdot (v_z, v_\lambda) = 0, \quad \forall \quad v_z \in T_z \mathcal{C}^{HP}, \quad v_\lambda \in T_\lambda C^0([a, b], \mathbb{R}^k).$$

Mean-square symplecticity of the flows of stochastic HP equations, and hence, constrained stochastic HP equations (by the equivalence) is established in [4]. \square

Eliminating the Lagrange Multiplier One can eliminate the Lagrange multiplier in (10) by taking the second Stratonovich differential of the constraint $g(q) = 0$:

$$\frac{dg}{dt}(q) = \frac{\partial g}{\partial q} \cdot v = 0$$

and,

$$d\frac{dg}{dt} = \frac{\partial^2 g}{\partial^2 q}(v, v)dt + \frac{\partial g}{\partial q} \cdot dv = 0. \quad (13)$$

By differentiating the Legendre transform in (10) we obtain,

$$dv = \left(\frac{\partial^2 L}{\partial^2 v} \right)^{-1} dp - \left(\frac{\partial^2 L}{\partial^2 v} \right)^{-1} \left(\frac{\partial^2 L}{\partial q \partial v} \right) v dt.$$

Substituting dp from (10) and the above into (13) gives,

$$\begin{aligned} \frac{\partial^2 g}{\partial^2 q}(v, v)dt + \frac{\partial g}{\partial q} \cdot \left(\frac{\partial^2 L}{\partial^2 v} \right)^{-1} \left(\frac{\partial L}{\partial q} dt + \sum_{i=1}^m \frac{\partial \gamma_i}{\partial q} dW_i + \frac{\partial g^*}{\partial q} d\lambda \right) \\ - \frac{\partial g}{\partial q} \cdot \left(\frac{\partial^2 L}{\partial^2 v} \right)^{-1} \left(\frac{\partial^2 L}{\partial q \partial v} \right) v dt = 0. \end{aligned}$$

Set $\mathbf{P}(q, v) = \frac{\partial g}{\partial q}(q) \left(\frac{\partial^2 L}{\partial^2 v}(q, v) \right)^{-1} \frac{\partial g}{\partial q}(q)^*$ and solve for $d\lambda$ in the above to obtain:

$$\begin{aligned} d\lambda = -\mathbf{P}(q, v)^{-1} \frac{\partial^2 g}{\partial^2 q}(v, v)dt - \mathbf{P}(q, v)^{-1} \frac{\partial g}{\partial q} \cdot \left(\frac{\partial^2 L}{\partial^2 v} \right)^{-1} \left(\frac{\partial L}{\partial q} dt + \sum_{i=1}^m \frac{\partial \gamma_i}{\partial q} dW_i \right) \\ + \mathbf{P}(q, v)^{-1} \frac{\partial g}{\partial q}(q) \cdot \left(\frac{\partial^2 L}{\partial^2 v} \right)^{-1} \left(\frac{\partial^2 L}{\partial q \partial v}(q, v) \right) v dt. \end{aligned}$$

Set $\mathbf{B}(q, v) = \frac{\partial g}{\partial q}(q) * \mathbf{P}(q, v)^{-1} \frac{\partial g}{\partial q}(q) \left(\frac{\partial^2 L}{\partial^2 v} \right)^{-1}$. Substitute $d\lambda$ into (10) gives,

$$\begin{cases} dq &= vdt, \\ dp &= (\mathbf{Id} - \mathbf{B}(q, v)) \left(\frac{\partial L}{\partial q}(q, v)dt + \sum_{j=1}^m \frac{\partial \gamma_j}{\partial q}(q) \circ dW_j \right) \\ &\quad - \frac{\partial g}{\partial q}(q) * \mathbf{P}(q, v)^{-1} \frac{\partial^2 g}{\partial^2 q}(v, v)dt + \mathbf{B}(q, v) \left(\frac{\partial^2 L}{\partial q \partial v}(q, v) \right) vdt, \\ p &= \frac{\partial L}{\partial v}(q, v). \end{cases} \quad (14)$$

Remark 3.2. One can check using properties of \mathbf{B} and \mathbf{P} , and the Stratonovich-Ito conversion formula that the Stratonovich correction term in (14) vanishes, and hence, (14) is equivalent to:

$$\begin{cases} dq &= vdt, \\ dp &= (\mathbf{Id} - \mathbf{B}(q, v)) \left(\frac{\partial L}{\partial q}(q, v)dt + \sum_{j=1}^m \frac{\partial \gamma_j}{\partial q}(q) dW_j \right) \\ &\quad - \frac{\partial g}{\partial q}(q) * \mathbf{P}(q, v)^{-1} \frac{\partial^2 g}{\partial^2 q}(v, v)dt + \mathbf{B}(q, v) \left(\frac{\partial^2 L}{\partial q \partial v}(q, v) \right) vdt, \\ p &= \frac{\partial L}{\partial v}(q, v). \end{cases} \quad (15)$$

4 Constrained, Stochastic VPRK Integrators

General Case Let $[a, b]$ and N be given, and define the fixed step size $h = (b - a)/N$ and $t_k = hk$, $k = 0, \dots, N$. The stochastic VPRK method will be derived by discretizing the kinematic constraint in the stochastic HP action integral with an s-stage RK approximant (cf. (3)), and the integral of the Lagrangian over the kth-time step by the following quadrature:

$$\int_{t_k}^{t_k+h} L(q, v)dt \approx hb_i L(Q_k^i, V_k^i).$$

Motivated by the methods introduced in [18; 19], the stochastic part of the stochastic HP action will be approximated by:

$$\int_{t_k}^{t_k+h} \gamma_r(q) \circ dW \approx (\nu_i \phi_r + \kappa_i \psi_r) \gamma_r(Q_k^i),$$

where $\phi_r : \Omega \rightarrow \mathbb{R}$ and $\psi_r : \Omega \rightarrow \mathbb{R}$ for $r = 1, \dots, m$. With this discretization 3/2-order strongly convergent methods can be derived. The constraint $g(q) = 0$ is enforced for all internal stage positions $\{Q_k^i\}_{i=1}^s$ using Lagrange multipliers as in the deterministic case.

Definition 4.1. Fix two points q_1 and q_2 on Q and define the **discrete constrained, stochastic VPRK action sum** $\mathfrak{G}_d : \Omega \times \mathcal{C}_d \rightarrow \mathbb{R}$ by:

$$\begin{aligned} \mathfrak{G}_d &= \sum_{k=0}^{N-1} \sum_{i=1}^s \left[hb_i L(Q_k^i, V_k^i) + \sum_{r=1}^m (\nu_i \phi_r + \kappa_i \psi_r) \gamma_r(Q_k^i) + hb_i \langle \Lambda_k^i, g(Q_k^i) \rangle \right. \\ &\quad \left. + h \left\langle p_k^i, (Q_k^i - q_k)/h - \sum_{j=1}^s a_{ij} V_k^j \right\rangle + h \left\langle p_{k+1}, (q_{k+1} - q_k)/h - \sum_{j=1}^s b_j V_k^j \right\rangle \right]. \end{aligned}$$

In the following we assume that the VPRK method has a well-defined solution much like the assumption made in theorem 2.5.

Theorem 4.2. *Given an s -stage RK method with b -vector and a -matrix, a Lagrangian system with smooth Lagrangian $L : TQ \rightarrow \mathbb{R}$ such that $\partial^2 L / \partial v^2$ is invertible, stochastic potential $\gamma_r : Q \rightarrow \mathbb{R}$ for $r = 1, \dots, m$, and smooth holonomic constraint function $g : Q \rightarrow \mathbb{R}^k$. A discrete curve $c_d \in \mathcal{C}_d(q_1, q_2)$ satisfies the following stochastic VPRK method:*

$$\left\{ \begin{array}{l} Q_k^i = q_k + h \sum_{j=1}^s a_{ij} V_k^j, \\ q_{k+1} = q_k + h \sum_{j=1}^s b_j V_k^j, \\ P_k^i = p_k + h \sum_{j=1}^s \left(b_j - \frac{b_j a_{ji}}{b_i} \right) \left(\frac{\partial L}{\partial q}(Q_k^j, V_k^j) + \frac{\partial g}{\partial q}(Q_k^j)^* \cdot \Lambda_k^j \right) \\ \quad + \sum_{j=1}^s \sum_{r=1}^m \left(1 - \frac{a_{ji}}{b_i} \right) (\nu_j \phi_r + \kappa_j \psi_r) \frac{\partial \gamma_r}{\partial q}(Q_k^j), \\ p_{k+1} = p_k + h \sum_{j=1}^s b_j \left(\frac{\partial L}{\partial q}(Q_k^j, V_k^j) + \frac{\partial g}{\partial q}(Q_k^j)^* \cdot \Lambda_k^j \right) \\ \quad + \sum_{j=1}^s \sum_{r=1}^m (\nu_j \phi_r + \kappa_j \psi_r) \frac{\partial \gamma_r}{\partial q}(Q_k^j), \\ P_k^i = \frac{\partial L}{\partial v}(Q_k^i, V_k^i), \\ g(Q_k^i) = 0. \end{array} \right. \quad (16)$$

for $i = 1, \dots, s$ and $k = 1, \dots, N-1$, if and only if it is a critical point of the function $\mathfrak{G}_d : \mathcal{C}_d \rightarrow \mathbb{R}$, that is, $\mathbf{d}\mathfrak{G}_d(c_d) = 0$. Moreover, there exist $\{\Lambda_k^i\}_{i=1}^s$ such that the discrete flow map defined by the above scheme, $F_h : T^*S \rightarrow T^*S$, preserves the canonical symplectic form on T^*S .

Proof. We assume the existence of a numerical solution of (16) and a discrete flow map $F_h : T^*S \rightarrow T^*S$.

The differential of $\mathfrak{G}_d(\omega, c_d)$ in the direction $z = (\{\delta q_k, \delta p_k\}, \{\delta Q_k^i, \delta V_k^i, p_k^i\}_{i=1}^s, \{\delta \Lambda_k^i\}_{i=1}^s)$ is given by:

$$\mathbf{d}\mathfrak{G}_d(\omega, c_d) \cdot z =$$

$$\begin{aligned} & \sum_{k=0}^{N-1} \sum_{i=1}^s h b_i \left[\frac{\partial L}{\partial q}(Q_k^i, V_k^i) \cdot \delta Q_k^i + \frac{\partial L}{\partial v}(Q_k^i, V_k^i) \cdot \delta V_k^i \right] + \sum_{r=1}^m (\nu_i \phi_r + \kappa_i \psi_r) \frac{\partial \gamma_r}{\partial q}(Q_k^i) \cdot \delta Q_k^i \\ & + h \left[\left\langle p_k^i, (\delta Q_k^i - \delta q_k)/h - \sum_{j=1}^s a_{ij} \delta V_k^j \right\rangle + \left\langle p_{k+1}, (\delta q_{k+1} - \delta q_k)/h - \sum_{j=1}^s b_j \delta V_k^j \right\rangle \right] \\ & + h \left[\left\langle \delta p_k^i, (Q_k^i - q_k)/h - \sum_{j=1}^s a_{ij} V_k^j \right\rangle + \left\langle \delta p_{k+1}, (q_{k+1} - q_k)/h - \sum_{j=1}^s b_j V_k^j \right\rangle \right] \\ & + h b_i \left[\left\langle \delta \Lambda_k^i, g(Q_k^i) \right\rangle + \left\langle \Lambda_k^i, \frac{\partial g}{\partial q}(Q_k^i) \cdot \delta Q_k^i \right\rangle \right]. \end{aligned}$$

Collecting terms with the same variations and summation by parts using the

boundary conditions $\delta q_0 = \delta q_N = 0$ gives,

$$\begin{aligned}
\mathbf{d}\mathfrak{G}_d(\omega, c_d) \cdot z = & \sum_{k=1}^{N-1} \sum_{i=1}^s \sum_{r=1}^m \left(h b_i \frac{\partial g}{\partial q}(Q_k^i)^* \Lambda_k^i + h b_i \frac{\partial L}{\partial q}(Q_k^i, V_k^i) + (\nu_i \phi_r + \kappa_i \psi_r) \frac{\partial \gamma_r}{\partial q}(Q_k^i) + p_k^i \right) \cdot \delta Q_k^i \\
& + \left(-p_{k+1} + p_k - \sum_{i=1}^s p_k^i \right) \cdot \delta q_k + h \left(b_i \frac{\partial L}{\partial v}(Q_k^i, V_k^i) - \sum_{j=1}^s a_{ji} p_k^j - b_i p_{k+1} \right) \cdot \delta V_k^i \\
& + h \left\langle \delta p_k^i, (Q_k^i - q_k)/h - \sum_{j=1}^s a_{ij} V_k^j \right\rangle + h \left\langle \delta p_{k+1}, (q_{k+1} - q_k)/h - \sum_{j=1}^s b_j V_k^j \right\rangle \\
& + h b_i \langle \delta \Lambda_k^i, g(Q_k^i) \rangle.
\end{aligned}$$

Since $\mathbf{d}\mathfrak{G}_d(\omega, c_d) = 0$ if and only if $\mathbf{d}\mathfrak{G}_d \cdot z = 0$ for all $z \in T_{c_d} \mathcal{C}_d$, one arrives at the desired equations with the elimination of p_k^i and the introduction of the internal stage variables $P_k^i = \partial L / \partial v(Q_k^i, V_k^i)$ for $i = 1, \dots, s$. Conversely, if c_d satisfies (16) then $\mathbf{d}\mathfrak{G}_d(\omega, c_d(\omega)) = 0$.

The proof of symplecticity using the variational principle is very similar to the proof of theorem 2.5 and is therefore omitted. \square

Constrained, Stochastic Variational Euler As an example we consider the following simple first-order strongly convergent integrator. One can derive higher-order accurate integrators for constrained systems following the procedure for unconstrained stochastic Hamiltonian systems described in [18; 19].

Let $B_r^k \sim \mathcal{N}(0, h)$ be normally distributed random variables for $r = 1, \dots, m$ and $k = 0, \dots, N-1$. The **constrained, stochastic variational Euler integrator** is given by:

$$q_{k+1} = q_k + h \hat{v}_{k+1}, \quad (17)$$

$$\hat{p}_{k+1} = p_k + h \frac{\partial L}{\partial q}(q_k, v_k) + \sum_{r=1}^m \frac{\partial \gamma_r}{\partial q}(q_k) B_r^k + h \frac{\partial g}{\partial q}(q_k)^* \Lambda_k^1, \quad (18)$$

$$0 = g(q_{k+1}), \quad (19)$$

$$\hat{p}_{k+1} = \frac{\partial L}{\partial v}(q_{k+1}, \hat{v}_{k+1}). \quad (20)$$

together with the projection step (7). Given (q_k, p_k) and h , this integrator determines (q_{k+1}, p_{k+1}) by eliminating \hat{v}_{k+1} using the Legendre transform, eliminating \hat{p}_{k+1} using (18), and determining Λ_k^1 by satisfying the constraint $g(q_{k+1}) = 0$. One then takes a projection step by solving (7) for p_{k+1} . The integrator defined by the composite map, $(q_k, p_k) \mapsto (q_{k+1}, \hat{p}_{k+1}) \mapsto (q_{k+1}, p_{k+1})$, has the following properties.

Theorem 4.3. *Constrained, stochastic variational euler is mean-squared symplectic and first-order strongly convergent integrator for (10).*

Proof. We will prove this by using the technique provided in [23].

Consider the following first-order strongly convergent, Euler-Maruyama integrator applied to (14):

$$\begin{cases} q_{k+1} &= q_k + hv_k, \\ p_{k+1} &= p_k + (\mathbf{Id} - \mathbf{B}(q_k, v_k)) \left(h \frac{\partial L}{\partial q}(q_k, v_k) + \sum_{j=1}^m \frac{\partial \gamma_j}{\partial q}(q_k) B_j^k \right) \\ &\quad - h \frac{\partial g}{\partial q}(q_k)^* \mathbf{P}(q_k, v_k)^{-1} \frac{\partial^2 g}{\partial^2 q}(v_k, v_k) + h \mathbf{B}(q_k, v_k) \left(\frac{\partial^2 L}{\partial q \partial v}(q_k, v_k) \cdot v_k \right), \\ p_{k+1} &= \frac{\partial L}{\partial v}(q_{k+1}, v_{k+1}). \end{cases} \quad (21)$$

The order of accuracy of constrained stochastic variational Euler will be determined by checking that the update determined by a single step of (17)-(20) with the projection step (7) agrees with a single step of (21) to $\mathcal{O}(h^2)$.

For this purpose the following expansions to $\mathcal{O}(h^2)$ are introduced:

$$\begin{aligned} q_{k+1} &= q_k + h\hat{v}_{k+1}, \quad \hat{v}_{k+1} = v_k + h\hat{v}_{k+1}^1, \quad \hat{p}_{k+1} = p_k + h\hat{p}_{k+1}^1 \\ v_{k+1} &= v_k + hv_{k+1}^1, \quad p_{k+1} = p_k + hp_{k+1}^1 \end{aligned}$$

Substituting these expansions into (17) confirms that the position update agrees to $\mathcal{O}(h^2)$. However, checking that the momentum or velocity updates agree to $\mathcal{O}(h^2)$ is more involved. To do this expand (19) in a Taylor series about q_k , and use $g(q_k) = 0$ and $\frac{\partial g}{\partial q}(q_k) \cdot v_k = 0$ to obtain,

$$g(q_{k+1}) = h^2 \frac{\partial g}{\partial q}(q_k) \cdot \hat{v}_{k+1}^1 + \frac{h^2}{2} \frac{\partial^2 g}{\partial^2 q}(q_k)(v_k, v_k) + \mathcal{O}(h^3) = 0 \quad (22)$$

From (18) it is clear that,

$$h\hat{p}_{k+1}^1 = h \frac{\partial L}{\partial q}(q_k, v_k) + \sum_{j=1}^m \frac{\partial \gamma_j}{\partial q}(q_k) B_j^k + h \frac{\partial g}{\partial q}(q_k)^* \Lambda_k^1. \quad (23)$$

Moreover the Taylor series expansion of (20) gives,

$$h \frac{\partial^2 L}{\partial^2 v}(q_k, v_k) \cdot \hat{v}_{k+1}^1 + h \frac{\partial^2 L}{\partial q \partial v}(q_k, v_k) \cdot v_k = h\hat{p}_{k+1}^1 + \mathcal{O}(h^2) \quad (24)$$

Using invertibility of $\frac{\partial^2 L}{\partial^2 v}(q_k, v_k)$, one can rewrite (24) as,

$$h\hat{v}_{k+1}^1 = -h \frac{\partial^2 L}{\partial^2 v}(q_k, v_k)^{-1} \frac{\partial^2 L}{\partial q \partial v}(q_k, v_k) \cdot v_k + h \frac{\partial^2 L}{\partial^2 v}(q_k, v_k)^{-1} \hat{p}_{k+1}^1 + \mathcal{O}(h^2)$$

Substitution of (23) into the above equation gives,

$$\begin{aligned} h\hat{v}_{k+1}^1 &= -h \left(\frac{\partial^2 L}{\partial^2 v} \right)_k^{-1} \left(\frac{\partial^2 L}{\partial q \partial v} \right)_k \cdot v_k \\ &\quad + \left(\frac{\partial^2 L}{\partial^2 v} \right)_k^{-1} \left(h \frac{\partial L}{\partial q}(q_k, v_k) + \sum_{j=1}^m \frac{\partial \gamma_j}{\partial q}(q_k) B_j^k + h \frac{\partial g}{\partial q}(q_k)^* \Lambda_k^1 \right) + \mathcal{O}(h^2) \end{aligned}$$

Substituting the above expression into (22) gives,

$$\begin{aligned}
& -h \left(\frac{\partial g}{\partial q} \right)_k \left(\frac{\partial^2 L}{\partial^2 v} \right)_k^{-1} \left(\frac{\partial^2 L}{\partial q \partial v} \right)_k \cdot v_k \\
& + \left(\frac{\partial g}{\partial q} \right)_k \left(\frac{\partial^2 L}{\partial^2 v} \right)_k^{-1} \left(h \frac{\partial L}{\partial q}(q_k, v_k) + \sum_{j=1}^m \frac{\partial \gamma_j}{\partial q}(q_k) B_j^k + h \frac{\partial g}{\partial q}(q_k)^* \Lambda_k^1 \right) \\
& + \frac{h}{2} \frac{\partial^2 g}{\partial^2 q}(q_k)(v_k, v_k) + \mathcal{O}(h^2) = 0
\end{aligned}$$

From which it follows that the Lagrange multiplier satisfies,

$$\begin{aligned}
h\Lambda_k^1 &= h\mathbf{P}_k^{-1} \left(\frac{\partial g}{\partial q} \right)_k \left(\frac{\partial^2 L}{\partial^2 v} \right)_k^{-1} \left(\frac{\partial^2 L}{\partial q \partial v} \right)_k \cdot v_k \\
& - \mathbf{P}_k^{-1} \left(\frac{\partial g}{\partial q} \right)_k \left(\frac{\partial^2 L}{\partial^2 v} \right)_k^{-1} \left(h \frac{\partial L}{\partial q}(q_k, v_k) + \sum_{j=1}^m \frac{\partial \gamma_j}{\partial q}(q_k) B_j^k \right) \\
& - \frac{h}{2} \mathbf{P}_k^{-1} \frac{\partial^2 g}{\partial^2 q}(q_k)(v_k, v_k) + \mathcal{O}(h^2)
\end{aligned}$$

Substitution into (23) gives,

$$\begin{aligned}
h\hat{p}_{k+1}^1 &= (\mathbf{Id} - \mathbf{B}_k) \left(h \frac{\partial L}{\partial q}(q_k, v_k) + \sum_{j=1}^m \frac{\partial \gamma_j}{\partial q}(q_k) B_j^k \right) \\
& + h\mathbf{B}_k \left(\frac{\partial^2 L}{\partial q \partial v} \right)_k \cdot v_k - \frac{h}{2} \frac{\partial g}{\partial q}(q_k)^* \mathbf{P}_k^{-1} \frac{\partial^2 g}{\partial^2 q}(q_k)(v_k, v_k). \quad (25)
\end{aligned}$$

To determine p_{k+1}^1 we expand the hidden velocity constraint in (7) to obtain,

$$\frac{\partial g}{\partial q}(q_{k+1})v_{k+1} = h \frac{\partial g}{\partial q}(q_k) \cdot v_{k+1}^1 + h \frac{\partial^2 g}{\partial^2 q}(q_k)(v_k, v_k) + \mathcal{O}(h^2) = 0 \quad (26)$$

Expanding the momentum update in (7) gives,

$$hp_{k+1}^1 = h\hat{p}_{k+1}^1 + h \frac{\partial g}{\partial q}(q_k)^* \Lambda_k^2 + \mathcal{O}(h^2). \quad (27)$$

Expanding the Legendre transform in (7) gives,

$$hv_{k+1}^1 = h \left(\frac{\partial^2 L}{\partial^2 v} \right)_k^{-1} \hat{p}_{k+1}^1 + h \left(\frac{\partial^2 L}{\partial^2 v} \right)_k^{-1} \frac{\partial g}{\partial q}(q_k)^* \Lambda_k^2 - h \left(\frac{\partial^2 L}{\partial^2 v} \right)_k^{-1} \left(\frac{\partial^2 L}{\partial q \partial v} \right)_k \cdot v_k + \mathcal{O}(h^2)$$

Substitution of the above into (26) yields,

$$\begin{aligned}
& h \frac{\partial g}{\partial q}(q_k) \cdot \left(\frac{\partial^2 L}{\partial^2 v} \right)_k^{-1} \hat{p}_{k+1}^1 + h\mathbf{P}_k \Lambda_k^2 - h \frac{\partial g}{\partial q}(q_k) \cdot \left(\frac{\partial^2 L}{\partial^2 v} \right)_k^{-1} \left(\frac{\partial^2 L}{\partial q \partial v} \right)_k \cdot v_k \\
& + h \frac{\partial^2 g}{\partial^2 q}(q_k)(v_k, v_k) + \mathcal{O}(h^2) = 0
\end{aligned}$$

Solving the above for Λ_k^2 and substituting into (27) gives,

$$hp_{k+1}^1 = h(\mathbf{Id} - \mathbf{B}_k)\hat{p}_{k+1}^1 + h\mathbf{B}_k \left(\frac{\partial^2 L}{\partial q \partial v} \right)_k \cdot v_k - h \frac{\partial g}{\partial q}(q_k)^* \mathbf{P}_k^{-1} \frac{\partial^2 g}{\partial^2 q}(q_k)(v_k, v_k) + \mathcal{O}(h^2)$$

Substituting (25) into the above and using the following identities

$$(\mathbf{Id} - \mathbf{B}_k)\mathbf{B}_k = 0, \quad (\mathbf{Id} - \mathbf{B}_k) \frac{\partial g}{\partial q}(q_k)^* \mathbf{P}_k^{-1} \frac{\partial^2 g}{\partial^2 q}(q_k)(v_k, v_k) = 0$$

implies that,

$$\begin{aligned} p_{k+1} = p_k + (\mathbf{Id} - \mathbf{B}_k) & \left(h \frac{\partial L}{\partial q}(q_k, v_k) + \sum_{j=1}^m \frac{\partial \gamma_j}{\partial q}(q_k) B_j^k \right) \\ & + h\mathbf{B}_k \left(\frac{\partial^2 L}{\partial q \partial v} \right)_k \cdot v_k - h \frac{\partial g}{\partial q}(q_k)^* \mathbf{P}_k^{-1} \frac{\partial^2 g}{\partial^2 q}(q_k)(v_k, v_k) + \mathcal{O}(h^2) \end{aligned} \quad (28)$$

which agrees with (21) to $\mathcal{O}(h^2)$.

Mean-square symplecticity of the map $(q_k, p_k) \mapsto (q_{k+1}, \hat{p}_{k+1})$ follows from the proof of theorem 4.2. The projection step is also mean-square symplectic. Therefore the composite map is mean-square symplectic with respect to the symplectic form on T^*S since the space of symplectic maps forms a group. \square

Example 4.1 (Stochastically Perturbed Spherical Pendulum). *Consider the spherical pendulum where $Q = \mathbb{R}^3$ and $S = S^2$ with holonomic constraint given by the zero level-set of $g(q) = \|q\|^2 - 1$. Let $\{e_1, e_2, e_3\}$ denote an orthonormal basis for \mathbb{R}^3 with e_3 corresponding to the direction of gravity. Its Lagrangian is given by:*

$$L(q, v) = \frac{1}{2}\|v\|^2 - q \cdot e_3.$$

Consider the following stochastic potentials: $\gamma_i(q) = \sin(q \cdot e_i)$ for $i = 1, 2, 3$. Applying stochastic variational Euler to the resulting stochastically perturbed spherical pendulum one obtains the strong error plot shown in Figure 1 confirming theorem 4.3. Since a strong order of convergence implies at least the same weak order if not higher, it is not surprising that the weak error plot shown in Figure 2 seems to predict 3/2-order weak convergence. The dashed lines in the plots give appropriate reference slopes in each case.

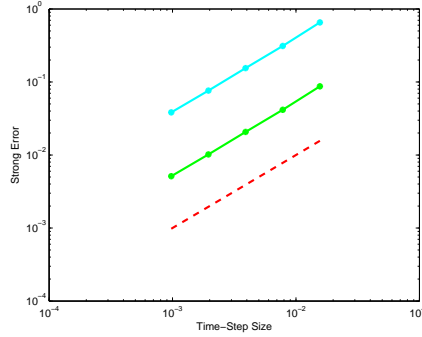


Figure 1: Strong error plot for stochastically perturbed spherical pendulum. A log-log graph of the strong error of stochastic variational Euler applied to the stochastically perturbed spherical pendulum. Green and cyan solid lines represent the strong errors in momentum and position respectively. For reference a dashed line of slope 1 is included. Observe that the order of convergence is consistent with theorem 4.3.

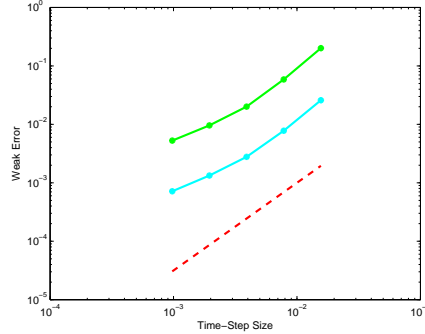


Figure 2: **Weak error plot for stochastically perturbed spherical pendulum.** A log-log graph of the weak error of stochastic variational Euler applied to the stochastically perturbed spherical pendulum. Green and cyan solid lines represent the strong errors in momentum and position respectively. For reference a dashed line with slope $3/2$ is included. This order of convergence is also consistent with theorem 4.3.

Acknowledgements

We wish to acknowledge Jerry Marsden, Sigrid Leyendecker, and Katie Whitehead for useful suggestions on this paper.

References

- [1] J. M. Bismut. *Mécanique aléatoire*. Springer, 1981.
- [2] N. Bou-Rabee. *Hamilton-Pontryagin Integrators on Lie Groups*. PhD thesis, California Institute of Technology, 2007.
- [3] N. Bou-Rabee and J. E. Marsden. Reduced hamilton-pontryagin variational integrators. Submitted; arXiv:, 2007.
- [4] N. Bou-Rabee and H. Owhadi. Stochastic variational integrators. Submitted; arXiv:0708.2187, 2007.
- [5] G. Ciccotti, T. Lelievre, and E. Vanden-Eijnden. Projections of diffusions on submanifolds: Application to mean force computation. *CPAM*, 61:0001–0039, 2008.
- [6] E. Hairer. Global modified hamiltonian for constrained symplectic integrators. *Numer. Math.*, 95:325–336, 2003.
- [7] E. Hairer, C. Lubich, and G. Wanner. *Geometric Numerical Integration*, volume 31 of *Springer Series in Computational Mathematics*. Springer, 2006.

- [8] E. Hairer, S. P. Nørsett, and G. Wanner. *Solving Ordinary Differential Equations I*. Springer-Verlag, 1993.
- [9] C. Hartmann. An ergodic sampling scheme for constrained hamiltonian systems with applications to molecular dynamics. Submitted to J. Stat. Phys., 2007.
- [10] L. Jay. Symplectic partitioned runge-kutta methods for constrained hamiltonian systems. *SIAM J. Num. Anal.*, 33:368–387, 1996.
- [11] J. A. Lazaro-Cami and J. P. Ortega. Reduction and reconstruction of symmetric stochastic differential equations. preprint; arXiv:0705.3156, 2007.
- [12] J. A. Lazaro-Cami and J. P. Ortega. Stochastic hamiltonian dynamical systems. preprint; math.PR/0702787, 2007.
- [13] B. Leimkuhler and R. Skeel. Symplectic numerical integrators in constrained hamiltonian systems. *JCP*, 112:117–125, 1994.
- [14] R. MacKay. *Some Aspects of the Dynamics of Hamiltonian Systems*, pages 137–193. Clarendon Press, 1992.
- [15] S. J. A. Malham and A. Wiese. Stochastic lie group integrators. *SISC*, 2007. To appear.
- [16] J. E. Marsden and M. West. Discrete mechanics and variational integrators. *Acta Numerica*, 10:357–514, 2001.
- [17] G. N. Milstein, Yu. M. Repin, and M. V. Tretyakov. Symplectic methods for hamiltonian systems with additive noise. *SIAM J. Num. Anal.*, 39:1–9, 2002.
- [18] G. N. Milstein, Yu. M. Repin, and M. V. Tretyakov. Symplectic methods for stochastic systems preserving symplectic structure. *SIAM J. Num. Anal.*, 40:1–9, 2003.
- [19] G. N. Milstein and M. V. Tretyakov. *Stochastic Numerics for Mathematical Physics*. Springer, 2004.
- [20] S. Reich. On higher-order semi-explicit symplectic partitioned runge-kutta methods for constrained hamiltonian systems. *Numer. Math.*, 76:231–247, 1997.
- [21] J. Ryckaert, G. Ciccotti, and H. Berendsen. Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes. *JCP*, 23:327–341, 1977.
- [22] Y. B. Suris. Hamiltonian methods of runge-kutta type and their variational interpretation. *Math. Modelling*, 2:78–87, 1990.

- [23] E. Vanden-Eijnden and G. Ciccotti. Second-order integrators for langevin equations with holonomic constraints. *Chem. Phys. Letters*, 429:310–316, 2006.
- [24] M. Veselov. Integrable discrete-time systems and difference operators. *Func. An. and Appl.*, 22:83–94, 1988.
- [25] J. M. Wendlandt and J. E. Marsden. Mechanical integrators derived from a discrete variational principle. *Physica D*, 106:223–246, 1997.
- [26] H. Yoshimura and J. E. Marsden. Dirac structures and lagrangian mechanics part i: Implicit lagrangian systems. *J. Geom. and Physics*, 57:133–156, 2006.
- [27] H. Yoshimura and J. E. Marsden. Dirac structures and lagrangian mechanics part ii: Variational structures. *J. Geom. and Physics*, 57:209–250, 2006.